

## Test Your Tech

Blogging is:

- A. Someone's online journal.
- B. A Celtic dance with wooden shoes.
- C. How the Celtics keep the Knicks away from the ball.

1

## Test Your Tech

Blogging is:

- A. Someone's online journal.
- B. A Celtic dance with wooden shoes.
- C. How the Celtics keep the Knicks away from the ball.

2

## Announcements

- Project 2B due tonight at 10pm
  - I will join the 1:30 drop-in lab in MGH 430 right after lecture
  - CLUE Tutoring tonight at 7pm in MGH 058

D.A. Clements, UW Information School

3

*FIT 100– Fluency with Information Technology*

## A Table with a View

*Data Storage and Transfer with  
XML and Databases*

D.A. Clements

D.A. Clements, UW Information School

4

## Differences Between Tables and Databases

- When we think of databases, we often think of tables of information
- Comparing Tables
  - Database tables
    - Metadata tag identifying each of the data fields
  - Spreadsheet tables
    - Rely on position to keep the integrity of their data
  - HTML tables
    - Data as table entries with no unique identity at all
    - Concerned only with how to display the data, not with its meaning

16-5

D.A. Clements, UW Information School

## The Database Advantage

- Metadata is key advantage of databases over other systems recording data as tables
- Two of the most important roles in defining metadata
  - Identify the type of data with a unique tag
  - Define the relationships of the data

16-6

D.A. Clements, UW Information School

## XML: A Language for Metadata Tags

- Extensible Markup Language
  - Tagging scheme similar to XHTML
  - No standard tags to learn
    - Self-describing, think up the tags you need
  - Works well with browsers and Web-based applications
  - Use a simple text editor
  - XML tag names cannot contain spaces

16-7 D.A. Clements, UW Information School

## Extensible Markup Language

# XML

D.A. Clements, UW Information School 8

## An Example from Tahiti

- Area in km<sup>2</sup> for Tahiti & neighboring islands

```
<?xml version = "1.0" encoding="ISO-8859-1" ?>
<archipelago>
<island><iName>Tahiti</iName> <area>1048</area></island>
<island><iName>Moorea</iName> <area>130</area></island>
<island><iName>Maiao</iName> <area>9.5</area></island>
<island><iName>Mehetia</iName> <area>2.3</area></island>
<island><iName>Tetiaroa</iName> <area>12.8</area></island>
</archipelago>
```

Figure 16.1 XML file encoding data for the Windward Islands database. The first line states that the file contains XML tags.

16-9 D.A. Clements, UW Information School

## An Example from Tahiti (cont'd)

- First line
  - <?xml version="1.0" encoding="ISO-8859-1" ?>
- File should be ASCII text
- File extension should be .xml

16-10 D.A. Clements, UW Information School

## Table 16.1 Rules for writing XML.

Required first line	<?xml version="1.0" encoding="ISO-8859-1" ?> must appear on the first line, starting in the first position.
First tag	The first tag encountered is the root element, and it must enclose all of the file's content; it appears on the second or possibly third line.
Closing tags	All tags must be closed.
Element naming	Observe these rules: <ul style="list-style-type: none"> <li>• Names can contain letters, numbers, and underscore characters.</li> <li>• Names must not start with a number or punctuation character.</li> <li>• Names must not start with the letters xml (or XML, or Xml, etc.).</li> <li>• Names cannot contain spaces.</li> </ul>
Case sensitivity	Tags and attributes are case sensitive.
Proper nesting	All tags must be well-nested.
Attribute quoting	All attribute values must be quoted; paired single quotes (apostrophes) or paired double quotes are okay; use "dumb" quotes only; choose 'opposite' quotes to enclose quoted values.
White space	White space is preserved and converted to a single space.
Comments	XML comments have the form: <!-- comment. -->

D.A. Clements, UW Information School

## Expanding Use of XML

- Combine encodings of two archipelagos – the Windward and the Galapagos Islands
- Root element is the tag that encloses all of the content of the XML file
  - <archipelago> in Fig. 16.1
  - <geo\_feature> in Fig. 16.2
- Indenting for readability and structure

16-12 D.A. Clements, UW Information School

```

<?xml version = "1.0"
encoding="ISO-8859-1" ?>
<geo_feature>
  <archipelago>
    <a_name>Windward Islands
    </a_name>
    <island>
      <iName>Tahiti</iName>
      <area>1048</area>
    </island>
    <island>
      <iName>Moorea</iName>
      <area>130</area>
    </island>
    <island>
      <iName>Maiao</iName>
      <area>9.5</area>
    </island>
    <island>
      <iName>Mehetia</iName>
      <area>2.3</area>
    </island>
    <island>
      <iName>Tetiarioa</iName>
      <area>12.8</area>
    </island>
  </archipelago>
  <archipelago>
    <a_name>Galapagos Islands
    </a_name>
    <island>
      <iName>Isabella</iName>
      <area>4588</area>
      <elevation>1707</elevation>
    </island>
    <island>
      <iName>Fernandina</iName>
      <area>642</area>
      <elevation>1494</elevation>
    </island>
    <island>
      <iName>Tower</iName>
      <area>14</area>
      <elevation>76</elevation>
    </island>
    <island>
      <iName>Santa Cruz</iName>
      <area>986</area>
      <elevation>846</elevation>
    </island>
  </archipelago>
</geo_feature>

```

Figure 16.2 XML file for the Geographic Features database. XML ignores white space, so the text in the file has been indented for easier reading.

## Attributes in XML

- Use attributes for additional metadata, not for additional content
  - Not good, name is content:
 

```
<archipelago name="Galapagos">
```
  - Better to give alternate form of the data
 

```
<a_name accents="Gal&acute;pagos">Galapagos</a_name>
```

## Effective Design with XML Tags

- Identification Rule: Label Data with Tags Consistently
  - You can choose whatever tag names you wish to name data, but once you've decided on a tag for a particular kind of data, you must always surround it with that tag.

## Effective Design with XML Tags (cont'd)

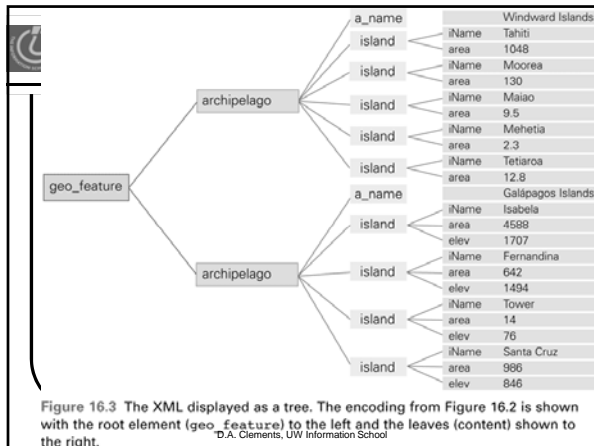
- Affinity Rule: Group Related Data
  - Enclose in a pair of tags all tagged data referring to the same entity. Grouping it keeps it all together, but the idea is much more fundamental: Grouping makes an association of the tagged data items as being related to each other, properties of the same thing.
  - Groups together data for a single thing – an island
    - Association is among properties of an object

## Effective Design with XML Tags (cont'd)

- Collection Rule: Group Related Instances
  - When you have several instances of the same kind of data, enclose them in tags; again, it keeps them together and implies that they are related by being instances of the same type.
  - Groups together data of several instance of the same thing – islands
    - Association is among the objects themselves (entities)

## The XML Tree

- XML encodings of information produce hierarchical descriptions that can be thought of as trees
  - Hierarchy a consequence of how tags enclose one another and the data



# DATABASES

D.A. Clements, UW Information School 20

## What is a Database

- Any organized collection of data
- A collection of similar data
- Examples of databases:
  - Telephone book white pages
  - T.V. Guide
  - Airline reservation system
  - Motor vehicle registration records

D.A. Clements, UW Information School 21

## Why do we need a database?

- Keep records of our:
  - Clients
  - Staff
  - Volunteers
- To keep a record of activities and interventions
- Keep sales records
- Develop reports
- Perform research

D.A. Clements, UW Information School 22

## Database Terminology

Phone book:	Fields (columns)	
Records (rows)	Anderson Thomas A 123 Marine View Dr. 237-1234	
	Benson Karen C 1300 California Ave 237-1098	
	Cassery Rick W 12492 Rd 19 342-0502	
	Drummond Lynn M 12059 30th Ave W 931-1105	
	Table	
Field (the columns in a table)	• Smallest unit of information in a table Sometime called "attributes"	• First name • Last name • Middle initial • Street address • Phone number(s)
Record (the rows in a table)	• All related fields are collectively called a record	• All fields for one person are a record
Table	• A collection of records is a data table	• Collection of everyone's records
Database Management System (DBMS)	• All the related tables, queries, data entry and edit forms, reports, macros and VBA modules constitute a database.	

D.A. Clements, UW Information School 23

## Database Management System (DBMS)

- Software tools for working with data
- Designed to:
  - Store (tables)
  - Organize (sort)
  - Add, modify or delete
  - Ask questions (queries)
  - Produce forms and reports
    - Summarizing
    - Displaying details
- Toolbox is a good analogy

D.A. Clements, UW Information School 24

## Ultimate Purpose of a Database Management System (DBMS)

To transform

```

    graph LR
    Data --> Information
    Information --> Knowledge
    Knowledge --> Action
  
```

D.A. Clements, UW Information School 25

## Flat-File vs. Relational Database

- Flat-File Database
  - All relevant data in a single table, or series of unrelated tables
  - Work best for small quantities of data; where viewing and sorting the data in a single list does not create a time-consuming task
  - Typically a person's first databases
  - Example: Excel spreadsheet or Word data list file
- Relational Database
  - Provide a solution to data entry redundancy problems
  - Linked through common fields (columns) with exactly the same data
  - Tables linked together can be queried as if one table
  - Can answer very complex questions

D.A. Clements, UW Information School 26

## Flat-File Example

### Staff Telephone List

Last Name	First Name	Emp ID	Dept	Location	Work Phone	M/S	Supervisor Name	Supr Phone
Adams	Wes	19589	PROD	Seattle	(206) 221-1958	QR-07	Susan Buckle	(206) 221-2241
Alberts	George	21533	PROD	Seattle	(206) 221-2153	QR-35	Marsha Mosley	(206) 221-1975
Allen	Susan	20256	PROD	Renton	(206) 393-2025	PB-18	Frank Sullivan	(206) 393-1000
Allert	Maria	10544	PROD	Seattle	(206) 221-1054	QR-27	Lynn Jarret	(206) 221-1366
Andrews	Mike	22113	PROD	Seattle	(206) 221-2211	QR-12	Harry Hillis	(206) 221-2179
Apperly	Ward	12244	PROD	Renton	(206) 393-1224	PB-14	Molly Goldberg	(206) 393-1513
Arthur	Diane	12370	MKTG	Bellevue	(206) 882-1237	RL-27	Wes Adams	(206) 221-1958
Asher	Jane	11222	ACCT	Seattle	(206) 221-1122	BX-45	Valarie Johnson	(206) 221-1958
Astor	Lawrence	20286	PROD	Seattle	(206) 221-2028	QR-10	Maggie Kramer	(206) 221-2028
Ayres	William	22263	PROD	Seattle	(206) 221-2226	QR-10	P. Kramer	(206) 221-2226
Baker	Gerald	19042	ACCT	Seattle	(206) 221-1904	BX-45	Valarie Johnson	(206) 221-1958

- Weaknesses common to flat-file systems
  - Duplicate information in the table
  - Inconsistencies in the way Supervisor Names are entered

D.A. Clements, UW Information School 27

## Relational Database Example

D.A. Clements, UW Information School 28

## Database Tables

D.A. Clements, UW Information School 29

## Query from Two Tables

D.A. Clements, UW Information School 30

## Forms

The screenshot displays two overlapping web forms. The top form, titled 'Employee Salary Data', includes fields for 'Last Name' (Adams), 'First Name' (Wes), 'Grade' (22), 'Years' (9.2), 'Monthly Salary' (\$3,812.55), 'Annual Salary' (\$45,750.60), and 'Hourly Rate' (\$21.91). The bottom form, titled 'XYZ Department Summary', shows 'Dept Code: ACCT', 'Building: A', and a summary of 20 employees with a total monthly salary of \$53,048.53. A small profile picture of Valerie Smith, Vice President, is visible in the bottom right corner of the summary form.

## Reports

The screenshot shows a 'Training Records by Employee' report. It features a table with columns for 'Emp ID', 'Emp Name', 'First Name', 'Course Name', 'Start Date', 'End Date', 'Hours', 'Status', 'Dept', 'Grade', and 'Exp'. The report lists training records for various employees, including Valerie Smith. A summary box in the top right corner provides department statistics for 'Accounting & Finance', including the number of employees (20), total monthly salary (\$53,048.53), and average monthly salary (\$2,652.43). A small profile picture of Valerie Smith is also present.

# RELATIONAL DATABASES

## Video

- Relational databases and tables